

# On the time constant of high dimensional first passage percolation

Antonio Auffinger <sup>\*</sup>

Northwestern University

Si Tang <sup>†</sup>

University of Chicago

January 29, 2016

## Abstract

We study the time constant  $\mu(e_1)$  in first passage percolation on  $\mathbb{Z}^d$  as a function of the dimension. We prove that if the passage times have finite mean,

$$\lim_{d \rightarrow \infty} \frac{\mu(e_1)d}{\log d} = \frac{1}{2a},$$

where  $a \in [0, \infty]$  is a constant that depends only on the behavior of the distribution of the passage times at 0. For the same class of distributions, we also prove that the limit shape is not an Euclidean ball, nor a  $d$ -dimensional cube or diamond, provided that  $d$  is large enough.

## 1 Introduction and main results

We study first passage percolation on  $\mathbb{Z}^d$  for  $d$  large. The model is defined as follows. We place a non-negative random variable  $\tau_e$ , called the passage time of the edge  $e$ , at each nearest-neighbor edge in  $\mathbb{Z}^d$ . The collection  $(\tau_e)$  is assumed to be independent, identically distributed with common distribution  $F$ .

A *path*  $\gamma$  is a finite or infinite sequence of nearest neighbor edges in  $\mathbb{Z}^d$  such that each two consecutive edges in the sequence intersect. For any finite path  $\gamma$  we define the passage time of  $\gamma$  to be

$$T(\gamma) = \sum_{e \in \gamma} \tau_e.$$

Given two points  $x, y \in \mathbb{Z}^d$  one then sets

$$T(x, y) = \inf_{\gamma} T(\gamma),$$

---

<sup>\*</sup>tuca@northwestern.edu. The research of A. A. is supported by NSF grant DMS-1597864.

<sup>†</sup>sitang@galton.uchicago.edu

<sup>0</sup>MSC2000: Primary 60K35, 82B43.

where the infimum is over all finite paths  $\gamma$  that start at the point  $x$  and end at  $y$ . For a review and the current state of the art of the model, we invite the readers to see the recent notes [1] or the classical paper of Kesten [7].

Here we focus on the large  $d$  behavior of the time constant and limit shape of the model. These are defined as follows. Let  $e_1, \dots, e_d$  be the coordinate vectors of  $\mathbb{Z}^d$ .

The time constant  $\mu(e_1) \in [0, \infty)$  is defined as

$$\mu(e_1) = \lim_{n \rightarrow \infty} \frac{T(0, ne_1)}{n} \quad \text{a.s. and in } L^1.$$

If  $\mathbb{E}\tau_e < \infty$ ,  $\mu(e_1)$  exists. See [1, Theorem 2.1] and the discussion therein.

For each  $t \geq 0$  let

$$B(t) = \{y \in \mathbb{R}^d : T(0, [y]) \leq t\},$$

where  $[y]$  is the unique point in  $\mathbb{Z}^d$  such that  $y \in [y] + [0, 1)^d$ . The pair  $(\mathbb{Z}^d, T(\cdot, \cdot))$  is a pseudo-metric space and  $B(t) \cap \mathbb{Z}^d$  is the (random) ball of radius  $t$  around the origin. The limit shape is defined by the famous shape theorem as follows.

Assume

$$\mathbb{E} \min\{t_1^d, \dots, t_{2d}^d\} < \infty, \tag{1.1}$$

where  $t_i, i = 1, \dots, 2d$ , are independent copies of  $\tau_e$  and

$$F(0) < p_c(d), \tag{1.2}$$

where  $p_c(d)$  is the threshold for bond percolation in  $\mathbb{Z}^d$ . We write  $rS = \{rs : s \in S\}$  for any subset  $S \subseteq \mathbb{R}^d$  and  $r \in \mathbb{R}$ .

**Theorem 1.1** (Cox and Durrett [3]). *If (1.1) and (1.2) hold, then the first passage percolation model has a limit shape. That is, there exists a deterministic, convex, compact set  $\mathcal{B}$  in  $\mathbb{R}^d$ , such that for each  $\varepsilon > 0$ ,*

$$\mathbb{P} \left( (1 - \varepsilon)\mathcal{B} \subset \frac{B(t)}{t} \subset (1 + \varepsilon)\mathcal{B} \text{ for all } t \text{ large} \right) = 1.$$

*Moreover, the limit shape  $\mathcal{B} = \{x \in \mathbb{R}^d : \mu(x) \leq 1\}$  has a non-empty interior and is symmetric about the axes of  $\mathbb{R}^d$ .*

Despite the importance of both objects, our knowledge on  $\mu(e_1)$  and  $\mathcal{B}$  is almost non-existent. Finding a distribution where one can explicitly determine  $\mu(e_1)$  is considered a difficult open task, and so is deriving further properties of  $\mathcal{B}$ . The main purpose of this paper is to investigate such questions for  $d$  large.

We assume that the passage times have finite mean,

$$\mathbb{E}\tau_e = \int_0^\infty x dF(x) < \infty \tag{1.3}$$

and the existence of some constant  $a \in [0, \infty]$  such that,

$$\left| \frac{\mathbb{P}(\tau_e \leq x)}{x} - a \right| \leq C \cdot |\log x|^{-1}, \tag{1.4}$$

for some  $C > 0$  in some interval  $[0, \varepsilon_0]$ ,  $\varepsilon_0 > 0$ . Here, we understand  $a = \infty$  as

$$\lim_{x \rightarrow 0} \frac{\mathbb{P}(\tau_e \leq x)}{x} = \infty. \quad (1.5)$$

Our first main result is the asymptotic behavior of  $\mu(e_1)$  as a function of  $d$ .

**Theorem 1.2.** *Assume (1.3) and that (1.4) holds for some  $a \in [0, \infty]$ . Then the time constant satisfies*

$$\lim_{d \rightarrow \infty} \frac{\mu(e_1)d}{\log d} = \frac{1}{2a}. \quad (1.6)$$

We now put the theorem above into historical context. The behavior of the time constant  $\mu(e_1)$  as a function of  $d$  was considered before by Kesten [7, Section 8] and Dhar [4]. Their assumptions on the distribution of the passage times are special cases of ours. First, in [7], under (1.3), (1.4) and the additional assumptions that  $a \in (0, \infty)$ ,  $C = o(1)$  as  $x \rightarrow 0$  and  $\tau_e$  has a density around the origin, Kesten showed the existence of  $\varepsilon > 0$  so that

$$\frac{\varepsilon}{a} < \liminf_{d \rightarrow \infty} \frac{\mu(e_1)d}{\log d} \leq \limsup_{d \rightarrow \infty} \frac{\mu(e_1)d}{\log d} \leq \frac{11}{a}. \quad (1.7)$$

Second, in [4], Dhar established (1.6) in the case of exponentially distributed passage times. Dhar's proof however cannot be adapted to any other distribution as, for instance, it heavily relies on the Markovian property of the ball  $B(t)$ . Thus, Theorem 1.2 says that the asymptotics obtained by Dhar are valid under rather general assumptions, that include those of Kesten.

*Remark 1.3.* Hypothesis (1.3) is a natural condition on the behavior of the distribution of the passage times at 0. It is satisfied by a large collection of examples; for instance, it includes all distributions that have a continuous density near the origin. The bound with  $|\log x|^{-1}$  is a weaker condition than any polynomial bound around 0.

*Remark 1.4.* If the distribution has a mass at 0 then (1.5) clearly holds. In this case, Theorem 6.1 in [7] and the fact that the critical probability  $p_c(d)$  for bond percolation in  $\mathbb{Z}^d$  decreases to 0 as  $d$  goes to infinity imply that the sequence  $(\mu(e_1))_{d \geq 1}$  will be eventually constant equal to 0 and, of course, (1.6) holds.

A word of comment is needed here. If, for some  $\delta > 0$ , the support of the distribution of the passage times is included in  $(\delta, \infty)$ , then it is clear that (1.6) must hold. Indeed, in this case,  $a = 0$  and  $\mu(e_1) \geq \delta$  for all  $d$ . It would be interesting to further study the behavior of  $\mu(e_1)$  in this situation. As we will see, this question seems to be related to the typical length (number of edges) of a geodesic and the behavior of  $p_c(d)$  as a function of  $d$ .

Our second main result excludes the  $d$ -dimensional Euclidean ball  $\mathbf{B} := \{x \in \mathbb{R}^d : \|x\|_2 \leq \mu(e_1)^{-1}\}$ , cube  $\mathbf{C} := \{x \in \mathbb{R}^d : \|x\|_\infty \leq \mu(e_1)^{-1}\}$  and diamond  $\mathbf{D} := \{x \in \mathbb{R}^d : \|x\|_1 \leq \mu(e_1)^{-1}\}$  as possible limit shapes. Note that due to convexity we always have  $\mathbf{D} \subseteq \mathbf{B} \subseteq \mathbf{C}$ .

**Theorem 1.5.** *For any distribution satisfying (1.1) and (1.2) for all  $d \geq 2$ , and (1.3), (1.4) with  $a \in (0, \infty)$ , there exists  $d_0 \geq 1$  such that for any  $d \geq d_0$ ,*

$$\mathbf{D} \subsetneq \mathbf{B} \subsetneq \mathbf{C} \text{ and } \mathbf{B} \neq \mathbf{B}.$$

*Remark 1.6.* We prove Theorem 1.5 by showing that the intersection of  $\mathcal{B}$  with the line  $\ell_d := \{\lambda(1, \dots, 1) : \lambda \in \mathbb{R}\}$  is strictly contained in  $\mathcal{B}$  and strictly contains  $\mathcal{D} \cap \ell_d$ . Using symmetry around  $\ell_d$  the proof may further exclude other possible limit shapes. We delay the proof until Section 5.

*Remark 1.7.* One of the main features of the Theorem above is that  $d_0$  can be explicitly estimated for any given values of  $a$  and  $C$  in (1.4). In the case of an exponential random variable or a uniform random variable on some interval  $[0, s]$ , we show that  $d_0 = 269,000$  is sufficient (but certainly not optimal). We exclude the  $d$ -dimensional diamond for all  $d \geq 110$ . See Appendix.

The fact that limit shape is not an Euclidean ball is expected to hold for all  $d \geq 2$ . Kesten provided the first results in this direction. He showed (see [7, Remark 8.5]) that this is the case for the exponential distribution if  $d \geq 10^6$ .

*Remark 1.8.* In [2], Couronné, Enriquez and Gerin also considered FPP with exponential distributed passage times. They provided a constructive way to find an upper bound of order  $\log d/d$  for  $\mu(e_1)$ . They also claimed that the limit shape is not an Euclidean ball if  $d \geq 35$ . However, the argument presented in [2, Corollary 4] seems unclear to us. Their claim is obtained using numerical results provided in Table 1 of Dhar [4]. It is unclear whether the inequality  $\mu(35) \leq 0.93 \frac{\log 2d}{2d}$  appearing in [2] implies  $\mu(d) \leq C(d) \frac{\log 2d}{2d}$  for some  $C(d)$  that leads to the result for  $d > 35$ . In view of (1.6), the constant  $C(d)$  must approach 1 as  $d \rightarrow \infty$ , even if the numbers appearing in Table 1 in [4] are monotonically decreasing.

The rest of the paper is organized as follows. In Section 3, we will sketch the proof of Theorem 1.2. The two following sections are devoted to prove the bounds  $\limsup_{d \rightarrow \infty} \frac{\mu(e_1)ad}{\log d} \leq 1/2$  and  $\liminf_{d \rightarrow \infty} \frac{\mu(e_1)ad}{\log d} \geq 1/2$ , respectively. In Section 5, we prove Theorem 1.5 by deriving a lower bound for the time constant in the diagonal direction. In the last section, we provide the quantitative bounds to control  $d_0$ . Throughout the paper, we use  $e_1, e_2, \dots, e_d$  to denote the canonical base vectors of  $\mathbb{Z}^d$ .

## Acknowledgments

Both authors thank the American Institute of Mathematics for their support during the workshop “First passage percolation and related models”, in August 2015, where this project was initiated. Si Tang thanks the Department of Mathematics at Northwestern University for the hospitality during her visits.

## 2 Proof Strategy of Theorem 1.2

The proof strategy is motivated by [7]. The reader will see that the upper bound

$$\limsup_{d \rightarrow \infty} \frac{\mu(e_1)ad}{\log d} \leq 1/2 \tag{2.1}$$

contains the most technical part and the main new ideas of the paper. The lower bound

$$\liminf_{d \rightarrow \infty} \frac{\mu(e_1)ad}{\log d} \geq 1/2$$

follows closely from [7] and is less intricate. In this section we explain how to derive (2.1) and the differences between our approach and Kesten's proof of (1.7).

First, it is known (see for instance [7, pp.246]) that

$$\mu(e_1) \leq \mathbb{E}\tilde{s}_{0,1}$$

where

$$\tilde{s}_{0,n} := \inf \left\{ \begin{array}{l} T(\gamma) : \gamma \text{ is a path from } (0, 0, \dots, 0) \text{ to some point in } H_n \\ \text{which, except for its } \textit{final} \text{ point, is contained in } [0, 1) \times \mathbb{R}^{d-1} \end{array} \right\}$$

$$H_n := \{(x_1, \dots, x_d) \in \mathbb{Z}^d : x_1 = n\}.$$

We will derive the upper bound for  $\mu(e_1)$  by bounding  $\mathbb{E}\tilde{s}_{0,1}$  from above. The idea is to look for a path of length  $n = n(d)$  to  $H_1$  that has a very small passage time, and it is contained in a subspace  $\mathcal{H}$  of dimension  $p = p(d) < d$ . To prove the existence of such a favorable path, the second moment method is a natural approach. In [7], Kesten took  $p = \lfloor d/2 \rfloor$  and considered directed paths whose first  $(n-1)$  steps go along the positive directions  $+e_2, \dots, +e_{p+1}$  and then, at the last step, take the  $e_1$  direction to reach  $H_1$ . As a result, these paths are necessarily self-avoiding, which allows an estimate of the passage time using the sum of i.i.d. random variables. Since these paths only take the positive directions of  $+e_2, \dots, +e_{p+1}$ , no more than  $\lfloor d/2 \rfloor^{(n-1)}$  paths were considered.

Kesten's proof was then a trade-off between examining a large collection of paths and being able to estimate  $\tilde{s}_{0,1}$  using sum of i.i.d. random variables. His strategy led to the upper bound in (1.7). In order to get an optimal upper bound, we explore a subspace that is almost as large as the entire  $\mathbb{Z}^d$  by choosing  $p = d - o(d)$ . Furthermore, we allow paths to go along any of the  $2p$  possible directions,  $\pm e_2, \dots, \pm e_{p+1}$ . Under such setting, we are able to examine nearly all of the paths leading to  $H_1$  and obtain the optimal upper bound. The price to pay is that now some paths will be self-intersecting and thus we can not approximate their passage times by a sum of i.i.d.'s. Furthermore, the computation in the second moment method becomes elaborate. In the end of the day, the price is affordable as in high dimension, the majority of random walk paths are self-avoiding. The main estimation is done by carefully counting patterns of overlapping segments for a given pair of random walks in  $\mathbb{Z}^d$ . This main step is done in Section 3.4.

## 3 Proof of (2.1)

### 3.1 Setup

We are interested in the self-avoiding paths of length  $n$  from 0 to  $H_1$  whose first  $(n-1)$  steps use directions  $\pm e_2, \dots, \pm e_{p+1}$  and the last step is  $e_1$ . Denote by  $\mathcal{P}_n$  the set of all such paths. For  $\gamma \in \mathcal{P}_n$ , we write it as

$$\gamma = (S_0 = 0, S_1, S_2, \dots, S_n),$$

where  $S_k \in \mathbb{Z}^d$  such that (i)  $S_i \neq S_j$  whenever  $i \neq j$ , (ii)  $S_k - S_{k-1} \in \{\pm e_2, \dots, \pm e_{p+1}\}$  for  $1 \leq k \leq n-1$ , and (iii)  $S_n - S_{n-1} = e_1$ . Let  $N_{n,x}$  be the number of paths  $\gamma \in \mathcal{P}_n$  such that

$T(\gamma) \leq x$ . We choose

$$n = \lfloor \log d \rfloor, \quad x = \frac{\log d}{2(1-\delta)ad}, \quad p = d - \left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor \quad (3.1)$$

for some  $\delta, \eta > 0$  fixed, but we will eventually send  $\delta$  to 0.

### 3.2 First moment of $N_{n,x}$

By definition,  $\mathbb{E}N_{n,x}$  can be written as

$$\mathbb{E}N_{n,x} = \sum_{\gamma \in \mathcal{P}_n} \mathbb{P}(T(\gamma) \leq x) = |\mathcal{P}_n| \cdot \mathbb{P}(S_n \leq x)$$

We will need the following two lemmas to estimate  $|\mathcal{P}_n|$  and  $\mathbb{P}(S_n \leq x)$ . The first one is about the number of self-avoiding walks, for which the estimate has been improved over the years [5, 6, 8].

**Lemma 3.1.** *Let  $C_{n,d}$  denote the number of  $d$ -dimensional self-avoiding walks of length  $n$ .*

(a)  $\xi_d := \lim_{n \rightarrow \infty} C_{n,d}^{1/n}$  exists and  $\xi_d \geq 2d - 1 - \log(2d - 1)$  for all  $d \geq 1$ .

(b) As  $d \rightarrow \infty$ ,  $\xi_d$  has the following expansion

$$\xi_d = 2d - 1 - \frac{1}{2d} - \frac{3}{(2d)^2} + O\left(\frac{1}{(2d)^3}\right).$$

**Lemma 3.2.** *Let  $X_1, X_2, \dots$ , be i.i.d. nonnegative random variables, satisfying (1.4). Let  $S_n := \sum_{i=1}^n X_i$  be the partial sum. Then, for all  $n \geq 1$  and  $0 \leq x \leq \varepsilon_0$ , there is*

$$\frac{(ax)^n}{n!} (1 - C|\log x|^{-1})^n \leq \mathbb{P}(S_n \leq x) \leq \frac{(ax)^n}{n!} (1 + C|\log x|^{-1})^n$$

*Proof.* The result follows from (1.4) and a similar calculation as in [7, Lemma 8.8]. □

By Lemma 3.1(a) and sub-additivity [8, pp.9], we know that

$$[2p - 1 - \log(2p - 1)]^{n-1} \leq \xi_p^{n-1} \leq |\mathcal{P}_n| \leq (2p)^{n-1}.$$

Also, Lemma 3.2 and Stirling's formula imply

$$\frac{(axe)^n (1 - C|\log x|^{-1})^n}{n^n e \sqrt{n}} \leq \mathbb{P}(S_n \leq x) \leq \frac{(axe)^n (1 + C|\log x|^{-1})^n}{n^n \sqrt{2\pi n}}.$$

Putting them together, we get

$$\frac{[2p - 1 - \log(2p - 1)]^{n-1} (axe)^n (1 - C|\log x|^{-1})^n}{n^n e \sqrt{n}} \leq \mathbb{E}N_{n,x} \leq \frac{(2paxe)^n (1 + C|\log x|^{-1})^n}{2pn^n \sqrt{2\pi n}}.$$

We look at the left side. Firstly, note that  $x \rightarrow 0$  as  $d \rightarrow \infty$ . In particular,

$$|\log x| = \log(2(1 - \delta)a) + \log d - \log \log d = \log d(1 + o(1)).$$

Hence, with  $n = \lfloor \log d \rfloor$  and for all  $d$  large, we have

$$(1 - C|\log x|^{-1})^n = e^{n \log(1 - C|\log x|^{-1})} = e^{-Cn|\log x|^{-1} + O(n|\log x|^{-2})} > e^{-2C}.$$

This gives a lower bound for  $\mathbb{E}N_{n,x}$  for all  $d$  sufficiently large:

$$\begin{aligned} \mathbb{E}N_{n,x} &\geq \frac{e^{-2C} \left( \frac{2p-1-\log(2p-1)}{2d(1-\delta)} \right)^{\log d}}{\frac{2p-1-\log(2p-1)}{d} e \sqrt{\log d}} \geq \frac{e^{-2C} d^{\log \frac{2p-1-\log(2p-1)}{2d(1-\delta)}}}{2e \sqrt{\log d}} \\ &\geq \frac{e^{-2C} d^{\log \frac{1-\frac{1}{d} \left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor - \frac{1+\log(2d)}{2d}}}{1-\delta}}}{2e \sqrt{\log d}}. \end{aligned} \quad (3.2)$$

For  $\delta \in (0, 1)$  fixed and  $d$  large, we will have  $\delta > \frac{1}{d} \left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor + \frac{1+\log(2d)}{2d}$ , which implies  $\frac{1-\frac{1}{d} \left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor - \frac{1+\log(2d)}{2d}}{1-\delta} > 1$ . Hence,  $\mathbb{E}N_{n,x} \rightarrow \infty$  as  $d \rightarrow \infty$ .

### 3.3 Proof of the Upper Bound

By definition, the second moment of  $N_{n,x}$  can be written as

$$\mathbb{E}N_{n,x}^2 = \sum_{\gamma, \gamma' \in \mathcal{P}_n} \mathbb{P}(T(\gamma) \leq x, T(\gamma') \leq x).$$

Suppose for now that we are able to show that for some  $0 < A < \infty$ , there is

$$\mathbb{E}N_{n,x}^2 \leq A(\mathbb{E}N_{n,x})^2 \quad (3.3)$$

for all  $d$  large. By Cauchy-Schwarz inequality, we know

$$\mathbb{P}(N_{n,x} \geq 1) \geq \frac{(\mathbb{E}N_{n,x})^2}{\mathbb{E}N_{n,x}^2} \geq \frac{1}{A} > 0.$$

This means with positive probability, we can find a path  $\gamma \in \mathcal{P}_n$  from 0 to  $H_1$  such that  $T(\gamma) < x$ . The proof of (3.3) will be given in Section 2.5.

Let  $\mathcal{H}$  be the subspace spanned by  $\pm e_2, \dots, \pm e_{p+1}$ . Now we focus on the coordinates  $e_{p+2}, e_{p+3}, \dots, e_d$ . For  $p+2 \leq j \leq d$ , let  $E_j$  be the event that there exists a path from  $e_j$  to  $H_1$  such that, except for its final point, is contained in  $[0, 1) \times \mathbb{R}^{d-1} \cap (\mathcal{H} + e_j)$  and has  $T(\gamma) \leq x$ . By translation invariance and (3.3), we have

$$\mathbb{P}(E_j) \geq 1/A > 0.$$

Choose  $y = \frac{\delta \log d}{ad}$ . For  $p+2 \leq j \leq d$ , let  $F_j$  be the event  $\{\tau_{e_j} \leq y\} \cap E_j$ . By Lemma 3.2,

$$\mathbb{P}(F_j) = \mathbb{P}(\tau_{e_j} \leq y) \mathbb{P}(E_j) \geq \frac{1}{A} \mathbb{P}(\tau_{e_j} < y) \geq \frac{1}{A} ay(1 - C|\log y|^{-1}) \geq \frac{\delta \log d}{Ad}(1 + o(1)).$$

Furthermore, these  $F_j$ 's are independent, and if any of the events  $F_j$  happens, we will have

$$\tilde{s}_{0,1} \leq y + x = \frac{\log d}{2ad} \left( 2\delta + \frac{1}{1-\delta} \right).$$

Therefore,

$$\begin{aligned} \mathbb{E}\tilde{s}_{0,1} &\leq (y+x)\mathbb{E}\mathbf{1}_{\bigcup_{j=p+2}^d F_j} + \mathbb{E}\tau_{e_1}\mathbb{E}\mathbf{1}_{\bigcap_{j=p+2}^d F_j^c} \\ &\leq \frac{\log d}{2ad} \left( 2\delta + \frac{1}{1-\delta} \right) + \left( 1 - \frac{\delta \log d}{Ad}(1+o(1)) \right)^{d-p-1} \mathbb{E}\tau_{e_1}. \end{aligned} \quad (3.4)$$

Notice that, for all  $d$  sufficiently large,

$$\left( 1 - \frac{\delta \log d}{Ad}(1+o(1)) \right)^{d-p-1} \leq 2 \left( 1 - \frac{\delta \log d}{Ad}(1+o(1)) \right)^{\left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor} \leq 4e^{-\frac{1}{A\delta^\eta}}.$$

Thus, for any  $\eta > 0$ , the second term in (3.4) vanishes as  $\delta \rightarrow 0$ . This gives us  $\mathbb{E}\tilde{s}_{0,1} \leq \frac{\log d}{2ad}$  as desired.

### 3.4 Second moment of $N_{n,x}$

We are going to prove (3.3) in this section. We first rewrite the second moment according to the number  $l \leq n$  of overlapping edges between  $\gamma$  and  $\gamma'$ :

$$\mathbb{E}N_{n,x}^2 = \sum_{l=0}^n \sum_{\gamma, \gamma' \in \mathcal{P}_n} \mathbb{P}(T(\gamma) \leq x, T(\gamma') < x) \mathbf{1}_{\{|\gamma \cap \gamma'|=l\}}. \quad (3.5)$$

Note that since we only consider  $\gamma, \gamma' \in \mathcal{P}_n$ , which are self-avoiding, the condition  $\{|\gamma \cap \gamma'| = l\}$  is defined with no ambiguity to be the number of edges in  $\gamma$  that also appear in  $\gamma'$  (or vice versa). In what follows, we always write

$$\gamma = (S_0 = 0, S_1, \dots, S_{n-1}, S_n), \quad \gamma' = (S'_0 = 0, S'_1, \dots, S'_{n-1}, S'_n). \quad (3.6)$$

When  $l = n$ , due to the fact that they both start from the origin and are self-avoiding, we know  $\gamma = \gamma'$ . In this case we have

$$\sum_{\gamma, \gamma' \in \mathcal{P}_n} \mathbb{P}(T(\gamma) \leq x, T(\gamma') < x) \mathbf{1}_{\{|\gamma \cap \gamma'|=n\}} = \mathbb{P}(T(\gamma) \leq x) \cdot |\mathcal{P}_n| = \mathbb{E}N_{n,x}.$$

When  $l = 0$ ,  $\gamma$  and  $\gamma'$  do not share any edges.

$$\sum_{\gamma, \gamma' \in \mathcal{P}_n} \mathbb{P}(T(\gamma) \leq x, T(\gamma') < x) \mathbf{1}_{\{|\gamma \cap \gamma'|=0\}} \leq \mathbb{P}(S_n \leq x)^2 \cdot |\mathcal{P}_n|^2 = (\mathbb{E}N_{n,x})^2.$$

For other  $1 \leq l \leq n-1$ , we can write,

$$\begin{aligned} \mathbb{P}(T(\gamma) \leq x, T(\gamma') < x) \mathbf{1}_{\{|\gamma \cap \gamma'|=l\}} &\leq \mathbb{P}(T(\gamma \setminus (\gamma \cap \gamma')) \leq x, T(\gamma') < x) \mathbf{1}_{\{|\gamma \cap \gamma'|=l\}} \\ &\leq \mathbb{P}(S_{n-l} \leq x) \mathbb{P}(S'_n < x) \mathbf{1}_{\{|\gamma \cap \gamma'|=l\}}. \end{aligned}$$



Lemma 3.2 implies that for all  $1 \leq l \leq n-1$  and  $d$  large,

$$\begin{aligned} \frac{\mathbb{P}(S_{n-l} \leq x)}{\mathbb{P}(S_n \leq x)} &\leq \frac{\frac{(ax)^{n-l}}{(n-l)!} (1+C|\log x|^{-1})^{n-l}}{\frac{(ax)^n}{n!} (1-C|\log x|^{-1})^n} \leq \left(\frac{n}{ax}\right)^l \left(1 + \frac{2C|\log x|^{-1}}{1-C|\log x|^{-1}}\right)^n \\ &\leq \left(\frac{n}{ax}\right)^l e^{n \log[1 + \frac{2C}{|\log x| - C}]} \leq \left(\frac{n}{ax}\right)^l e^{\frac{2nC}{|\log x| - C}} = \left(\frac{n}{ax}\right)^l e^{f_{a,C}(\delta, d)}, \end{aligned} \quad (3.7)$$

where we denote

$$f_{a,C}(\delta, d) := \frac{2nC}{|\log x| - C} = \frac{2C \log d}{\log[2(1-\delta)ad] - \log \log d - C}. \quad (3.8)$$

Note, for fixed  $a, C, \delta > 0$ ,  $f_{a,C}(\delta, d) \rightarrow 2C$  as  $d \rightarrow \infty$ . Hence, for sufficiently large  $d$ ,  $f_{a,C}(\delta, d) < 3C$ , and (3.5) becomes

$$\begin{aligned} \mathbb{E}N_{n,x}^2 &\leq \mathbb{E}N_{n,x} + (\mathbb{E}N_{n,x})^2 + e^{3C} \mathbb{P}^2(S_n \leq x) \sum_{l=1}^{n-1} \sum_{\gamma, \gamma' \in \mathcal{P}_n} \left(\frac{n}{ax}\right)^l \mathbf{1}_{\{|\gamma \cap \gamma'| = l\}} \\ &\leq \mathbb{E}N_{n,x} + (\mathbb{E}N_{n,x})^2 + e^{3C} \mathbb{P}^2(S_n \leq x) \sum_{l=1}^{n-1} \left(\frac{n}{ax}\right)^l \#\{(\gamma, \gamma') : \gamma, \gamma' \in \mathcal{P}_n, |\gamma \cap \gamma'| = l\} \\ &= \mathbb{E}N_{n,x} + (\mathbb{E}N_{n,x})^2 + e^{3C} [(2p)^{n-1} \mathbb{P}(S_n \leq x)]^2 \sum_{l=1}^{n-1} \left(\frac{n}{ax}\right)^l \frac{\#\{(\gamma, \gamma') : \gamma, \gamma' \in \mathcal{P}_n, |\gamma \cap \gamma'| = l\}}{(2p)^{2(n-1)}}. \end{aligned}$$

Note that  $\mathbb{E}N_{n,x} \rightarrow \infty$  and the front factor of the second term satisfies, for  $d$  large,

$$1 \leq \frac{(2p)^{n-1} \mathbb{P}(S_n \leq x)}{\mathbb{E}N_{n,x}} \leq \left(1 + \frac{1 + \log(2p-1)}{2p-1 - \log(2p-1)}\right)^{n-1} \leq e^{g_\eta(\delta, d)},$$

where we have denoted the exponent by

$$\begin{aligned} g_\eta(\delta, d) &= \frac{(n-1)(\log(2p-1) + 1)}{2p-1 - \log(2p-1)} \\ &\leq \frac{(\log d)^2 + \log 2 \log d}{2d(1 - 1/(\delta^{1+\eta} \log d)) - 1 - \log 2d}. \end{aligned} \quad (3.9)$$

For  $\delta, \eta > 0$  fixed, and  $d$  sufficiently large,  $g_\eta(\delta, d) \rightarrow 0$  as  $d \rightarrow \infty$ . Hence, we can choose  $d$  sufficiently large, such that  $e^{g_\eta(\delta, d)} < 2$ , which yields

$$\mathbb{E}N_{n,x}^2 \leq [\mathbb{E}N_{n,x}]^2 \left\{ 1 + 4e^{3C} \sum_{l=1}^{n-1} \left(\frac{n}{ax}\right)^l \frac{\#\{(\gamma, \gamma') : \gamma, \gamma' \in \mathcal{P}_n, |\gamma \cap \gamma'| = l\}}{(2p)^{2(n-1)}} + o(1) \right\}.$$

To proceed, we need the following proposition.

**Proposition 3.3.** *For each  $1 \leq l \leq n-1$  fixed and  $p$  sufficiently large,*

$$\frac{\#\{(\gamma, \gamma') : \gamma, \gamma' \in \mathcal{P}_n, |\gamma \cap \gamma'| = l\}}{(2p)^{2(n-1)}} \leq \left(\frac{1}{2p}\right)^l (1 + o(p^{-1/2})). \quad (3.10)$$

Given Proposition 3.3, we have

$$\begin{aligned}\mathbb{E}N_{n,x}^2 &\leq [\mathbb{E}N_{n,x}]^2 \left[ 2 + 8e^{3C} \sum_{l=0}^{n-1} \left( \frac{n}{2pax} \right)^l \right] \\ &\leq [\mathbb{E}N_{n,x}]^2 \left[ 2 + 8e^{8C} \sum_{l=0}^{\infty} \left( \frac{1-\delta}{1-1/(\delta^{1+\eta} \log d)} \right)^l \right].\end{aligned}$$

Again, for  $\delta > 0$  fixed, and  $d$  sufficiently large, we have  $\frac{1-\delta}{1-1/(\delta^{1+\eta} \log d)} < 1$ , hence the summation above converges, which gives (3.3). The proof for the upper bound is then complete once we show (3.10).

When  $\gamma, \gamma' \in \mathcal{P}_n$  and  $|\gamma \cap \gamma'| = l$ , there are two cases: (i) if  $S_{n-1} \neq S'_{n-1}$ , it is necessary that all  $l$  overlapping edges occur in the first  $(n-2)$  steps since both  $\gamma$  and  $\gamma'$  take the  $e_1$  direction at the last step; (ii) if  $S_{n-1} = S'_{n-1}$ , then  $\gamma$  and  $\gamma'$  share the last edge so that there are at most  $(l-1)$  overlapping edges in their first  $(n-1)$  steps. Observe that the denominator of (3.10) is just the number of all pairs of paths in  $\mathbb{Z}^p$  of length  $(n-1)$ , starting from the origin. Hence, if we put uniform measure on all pairs of simple random walk paths  $(\tilde{\gamma}, \tilde{\gamma}')$  in  $\mathbb{Z}^p$  starting from the origin and of length  $(n-1)$ , the left side of (3.10) is bounded by

$$\mathbb{P}\left(\{\tilde{\gamma}, \tilde{\gamma}' \text{ are self-avoiding}\} \cap \{|\tilde{\gamma} \cap \tilde{\gamma}'| = l\}\right) \quad \dots \quad \text{Case (i)} \quad (3.11)$$

$$+ \mathbb{P}\left(\{\tilde{\gamma}, \tilde{\gamma}' \text{ are self-avoiding}\} \cap \{|\tilde{\gamma} \cap \tilde{\gamma}'| = l-1\} \cap \{S_{n-1} = S'_{n-1}\}\right) \quad \dots \quad \text{Case (ii)} \quad (3.12)$$

Here  $\tilde{\gamma}$  and  $\tilde{\gamma}'$  can be thought as the first  $(n-1)$  steps of  $\gamma$  and  $\gamma'$ , respectively, and we abuse the notation by writing

$$\tilde{\gamma} = (S_0 = 0, S_1, \dots, S_{n-1}), \quad \tilde{\gamma}' = (S'_0 = 0, S'_1, \dots, S'_{n-1}). \quad (3.13)$$

We prove Case (i) and Case (ii) in Lemma 3.4 (i) and (ii), respectively. The proofs for both cases are very similar, based on counting the “bubbles” of two intersecting simple random walk paths. We explain in full detail the construction in Case (i), whereas for Case (ii), we just point out the difference.

Let  $\gamma$  and  $\gamma'$  be two paths sampled uniformly and independently from all simple random walk paths in  $\mathbb{Z}^p$ , starting from the origin, and of length  $n \leq \lfloor 10 \log p \rfloor$ . Note that the constant 10 takes into account that we are actually interested in paths of length  $\lfloor \log d \rfloor - 1$  and  $p \approx d$ . Such differences are negligible when  $d$  is large.

**Lemma 3.4.** *For all  $2 \leq n \leq \lfloor 10 \log p \rfloor$ ,  $1 \leq l \leq n-1$ , and sufficiently large  $p$ , we have*

$$(i) \quad \mathbb{P}(\{\gamma, \gamma' \text{ are self-avoiding}\} \cap \{|\gamma \cap \gamma'| = l\}) \leq \left(\frac{1}{2p}\right)^l (1 + o(p^{-1/2})),$$

$$(ii) \quad \mathbb{P}(\{\gamma, \gamma' \text{ are self-avoiding}\} \cap \{|\gamma \cap \gamma'| = l-1\} \cap \{S_n = S'_n\}) \leq \left(\frac{1}{2p}\right)^l o(p^{-1/2}).$$

Case (i) and (ii) follow from Lemma 3.4, when we replace  $n$  by  $(n - 1)$  and  $\gamma, \gamma'$  by  $\tilde{\gamma}, \tilde{\gamma}'$ , respectively.

*Proof of Lemma 3.4 (i).*

For each  $1 \leq K \leq l$ , let  $B_K$  denote the event that the  $l$  overlapping edges are clustered in  $K$  consecutive pieces. Let  $n_1, \dots, n_K$  be the lengths of these segments,  $m_1, \dots, m_K$  (resp.  $m'_1, \dots, m'_K$ ) be the indices of their starting points in  $\gamma$  (resp. in  $\gamma'$ ). For example, in the left part of Figure 1, we have  $l = 5$  and  $K = 2$ . For convenience, we also denote by  $A_l$  the event  $|\gamma \cap \gamma'| = l$  and  $G_\gamma$  (resp.  $G_{\gamma'}$ ) the event that  $\gamma$  (resp.  $\gamma'$ ) is self-avoiding. The original probability is just

$$\mathbb{P}(G_\gamma \cap G_{\gamma'} \cap A_l) = \sum_{K=1}^l \mathbb{P}(G_\gamma \cap G_{\gamma'} \cap A_l \cap B_K).$$

Note that, on the event  $G_\gamma \cap G_{\gamma'}$ , the number of segments  $K$  must be the same in both  $\gamma$  and  $\gamma'$ , i.e., the situations in the middle and on the right of Figure 1 cannot happen.

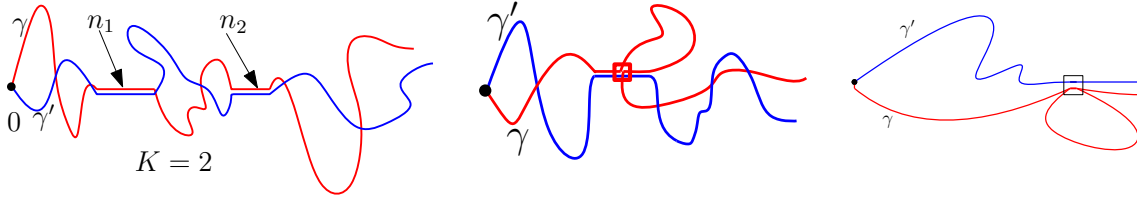


Figure 1: An illustration of the definition of  $K$  and  $n_1, \dots, n_K$ . Parallel segments in the figure represent overlapping edges. On the left, two overlapping segments with the same orientation. In the middle, two consecutive edges in  $\gamma$  overlap with two distant edges in  $\gamma'$ . On the right, two edges in  $\gamma$  overlap with the same edge in  $\gamma'$ . The last two situations cannot occur in the event  $G_\gamma \cap G_{\gamma'}$ .

When  $K = 1$ , the  $l$  overlapping edges are clustered in one segment. There are two possible alignments for the overlapping segment, either along the same or along different directions (see Figure 2). In both cases, nonoverlapping pieces of  $\gamma$  and  $\gamma'$  form a “bubble”-like shape.

The first situation happens with probability no more than

$$\begin{aligned} \sum_{m_1=0}^{n-l} \sum_{m'_1=0}^{n-l} \mathbb{P}(S_{m_1} = S'_{m'_1}) \left(\frac{1}{2p}\right)^l &= \left(\frac{1}{2p}\right)^l \left[ 1 + \sum_{m_1=0}^{n-l} \sum_{m'_1=0}^{n-l} \mathbb{P}(S_{m_1} = S'_{m'_1}) \mathbf{1}_{\{m_1+m'_1>0\}} \right] \\ &= \left(\frac{1}{2p}\right)^l \left[ 1 + \sum_{m_1=1}^{n-l} \sum_{m'_1=1}^{n-l} \mathbb{P}(S_{m_1+m'_1} = 0) \right] \\ &\leq \left(\frac{1}{2p}\right)^l \left[ 1 + \frac{(n-l)^2}{2p} \right]. \end{aligned}$$

We have used the fact that if  $\{S_m\}_{m \geq 0}$  is a simple random walk on  $\mathbb{Z}^p$ , then for all  $m \geq 1$  and  $t \in \mathbb{Z}^p$ ,

$$\mathbb{P}(S_m = t) \leq \sum_{s \in \mathbb{Z}^p} \mathbb{P}(S_{m-1} = s) \mathbb{P}(S_1 = t - s) \leq \frac{1}{2p} \sum_{s \in \mathbb{Z}^p} \mathbb{P}(S_{m-1} = s) \leq \frac{1}{2p}. \quad (3.14)$$

We write all terms for the second situation (Figure 2, right), as this will give us the spirit of the general  $K > 1$  case. The probability is no more than

$$\sum_{m_1, m'_1=0}^{n-l-1} \sum_{t, t', s, s' \in \mathbb{Z}^p} \mathbb{P} \left( \begin{array}{l} S_{m_1} = t, S_{l+m_1} - S_{m_1} = s, \\ S'_{m'_1} = t', S'_{l+m'_1} - S'_{m'_1} = s', \\ S'_{m'_1+j} = S_{m_1+l-j}, 0 \leq j \leq l \end{array} \right) \mathbf{1}_{\{m_1+m'_1 > 0\}}.$$

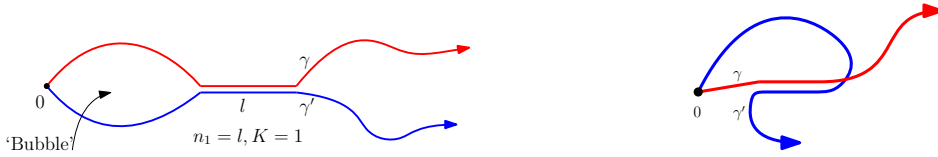


Figure 2: Case  $K = 1$ . The two paths  $\gamma$  and  $\gamma'$  intersect in a single segment and produce one bubble. They overlap along the same direction (left) or along opposite directions (right). Parallel segments in the figure represent overlapping edges.

Note that in this case, we can not have  $m_1 = m'_1 = 0$ , because otherwise the overlapping segments would have aligned along the opposite direction. Conditioning on the events  $\{S'_{m'_1} = t', S'_{l+m'_1} - S'_{m'_1} = s'\}$  and  $\{S'_{m'_1+j} = S_{m_1+l-j}, 0 \leq j \leq l\}$ , there is only one possibility for the choices of  $t$  and  $s$ , i.e.,  $s = -s'$ ,  $t = t' - s'$ . Moreover, the event  $S_{l+m_1} - S_{m_1} = -s'$  is implied by the event  $S'_{m'_1+j} = S_{m_1+l-j}$ ,  $0 \leq j \leq l$ . Thus, using the independence between  $\gamma$  and  $\gamma'$ , we can write

$$\begin{aligned} & \left( \frac{1}{2p} \right)^l \sum_{m_1, m'_1=0}^{n-l} \sum_{t', s' \in \mathbb{Z}^p} \mathbb{P}(S_{m_1} = t' - s') \mathbb{P}(S'_{m'_1} = t') \mathbb{P}(S'_{m'_1+l} - S'_{m'_1} = s') \mathbf{1}_{m_1+m'_1 > 0} \\ & \leq \left( \frac{1}{2p} \right)^l \sum_{m_1, m'_1=0}^{n-l} \sum_{t', s' \in \mathbb{Z}^p} \frac{1}{2p} \cdot \mathbb{P}(S'_{m'_1} = t') \mathbb{P}(S'_{m'_1+l} - S'_{m'_1} = s') \mathbf{1}_{m_1+m'_1 > 0} \\ & \leq \left( \frac{1}{2p} \right)^l \sum_{m_1, m'_1=0}^{n-l} \frac{1}{2p} \leq \left( \frac{1}{2p} \right)^l \frac{(n-l)^2}{2p}. \end{aligned}$$

Combining the two situations and using  $n = C_0 \log p$ , we conclude

$$\mathbb{P}(G_\gamma \cap G_{\gamma'} \cap A_l \cap B_1) \leq \left( \frac{1}{2p} \right)^l (1 + o(p^{-1/2})).$$

From the  $K = 1$  case, we have the following observations, and the last observation is the most important for the general  $1 < K \leq l$  case.

- a. The overlapping edges give the factor  $(1/2p)^l$ .

- b. There are either  $(K - 1)$  or  $K$  bubble segments (i.e., those that do not overlap with the other path) in  $\gamma$  and  $\gamma'$ , depending on whether  $m_1$  and  $m'_1$  are both zero or not. The total number of bubble and overlapping segments is either  $2K - 1$  or  $2K$ .
- c. Conditioning on the values of all  $2K - 1$  or  $2K$  segments of  $\gamma'$ , there is at most one solution for the values on the segments of  $\gamma$ . The event that “those  $K$  overlapping segments in  $\gamma$  take particular values” is absorbed in the event that “each edge in these segments overlaps with  $\gamma$ ”, whereas the event “the bubble segments take particular values” occurs with probability no more than  $\left(\frac{1}{2p}\right)^{K-1}$ , due to (3.14).

We now proceed to the general  $1 < K \leq l$  case. For each  $K$  fixed, we first compute the number of ways to divide the  $l$  overlapping edges into  $K$  (nonempty) groups of sizes  $n_1, n_2, \dots, n_K$ , which is no more than  $\binom{l-1}{K-1} \leq l^{K-1}$ . Next, we determine the positions of these  $K$  overlapping segments in  $\gamma$ : there are  $\binom{n}{K}$  ways to choose the starting points  $m_1 < m_2 < \dots < m_K$  of these segments, and once we have the starting points, we have  $K!$  ways to associate a running length  $n_{j_i}$  to a starting point  $m_i$ . This is over-counting, because if, say,  $m_2 - m_1 < n_3$ , then the overlapping segment starting at  $m_1$  can not be longer than  $n_3$ . We do the same for  $\gamma'$ . Once we have the locations of overlapping segments in  $\gamma$  and  $\gamma'$ , we know exactly which segment in  $\gamma$  overlap with which segments in  $\gamma'$ . There is an additional factor  $2^K$ , which counts for the two possible directions of alignment for each overlapping segment. These together give us a combinatorial number no larger than

$$l^{K-1} \cdot \left[ \binom{n}{K} \cdot K! \right]^2 2^K \leq \frac{(2ln^2)^K}{l},$$

which is an upper bound of all possible overlapping patterns (one of these is illustrated in Figure 3). Each pattern occurs with probability no more than

$$\left(\frac{1}{2p}\right)^l \left(\frac{1}{2p}\right)^{K-1},$$

and hence for all  $2 \leq K \leq l$ ,

$$\mathbb{P}(G_\gamma \cap G_{\gamma'} \cap A_l \cap B_K) \leq \left(\frac{1}{2p}\right)^l \left[ \frac{2l^{\frac{K-1}{K}} n^2}{(2p)^{\frac{K-1}{K}}} \right]^K \leq \left(\frac{1}{2p}\right)^l \left(\frac{ln^2}{p}\right)^K \frac{2p}{l}.$$

Summing over  $K$  and using the fact that  $K \leq l \leq n \leq C_0 \log p$ , we obtain

$$\frac{2p}{l} \sum_{K=2}^l \left(\frac{1}{2p}\right)^l \left(\frac{ln^2}{p}\right)^K = \left(\frac{1}{2p}\right)^l \frac{2ln^4}{p} \sum_{K=0}^{l-2} \left(\frac{ln^2}{p}\right)^K = \left(\frac{1}{2p}\right)^l \frac{2ln^4}{p} \frac{1}{1 - \frac{ln^2}{p}} = \left(\frac{1}{2p}\right)^l o(p^{-1/2}),$$

which finishes the proof of Lemma 3.4 (i).  $\square$

*Proof of Lemma 3.4 (ii).* When  $K \geq 3$ , there are at least  $(K - 1) \geq 2$  bubble segments in both  $\gamma$  and  $\gamma'$ . In this case, we can simply ignore the last event  $\{S_n = S'_n\}$  when calculating

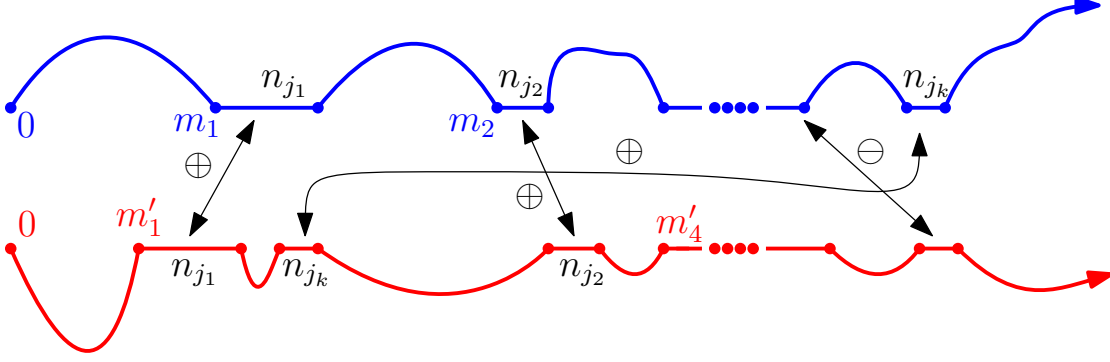


Figure 3: Overlapping paths  $\gamma$  and  $\gamma'$  and their matching: we split the  $l$  overlapping edges into groups of sizes  $n_1, n_2, \dots$ , starting at points  $m_1, m_2, \dots$ . Each overlapping piece can meet in a positive or negative orientation.

the probability. Following the same strategy as in Case (ii), one can easily get

$$\begin{aligned} \left(\frac{1}{2p}\right)^{l-1} \sum_{K=3}^l \frac{(2ln^2)^K}{l \cdot [2p]^{K-1}} &= \left(\frac{1}{2p}\right)^l \sum_{K=1}^{l-2} \frac{(2ln^2)^{K+2}}{l \cdot [2p]^K} = \frac{4l^2n^6}{p} \left(\frac{1}{2p}\right)^l \sum_{K=0}^{l-3} \left(\frac{ln^2}{p}\right)^K \\ &\leq \left(\frac{1}{2p}\right)^l \frac{4l^2n^6}{p} \frac{1}{1 - \frac{ln^2}{p}} = \left(\frac{1}{2p}\right)^l o(p^{-1/2}). \end{aligned}$$

When  $K = 1$  and if  $m_1 = m'_1 = 0$  (i.e.,  $\gamma$  and  $\gamma'$  overlap at the first  $l - 1$  edges) or  $m_1 = m'_1 = (n - l + 1)$  (i.e.,  $\gamma$  at  $\gamma'$  overlap at the last  $l - 1$  edges), there is only bubble segment in both  $\gamma$  and  $\gamma'$ . Also, on the event that they are both self-avoiding and  $S_0 = S'_0, S_n = S'_n$ , the overlapping edges must align in the same direction. This case is illustrated in Figure 4.

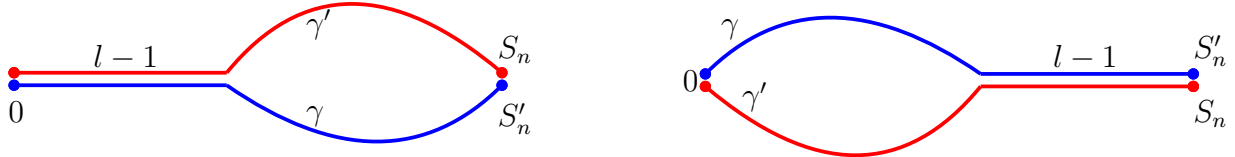


Figure 4: Representation of the case  $K = 1$  with an aligned overlap. Only one bubble happens either at the end or the beginning of the paths.

In either situation, the probability of seeing such a “bubble” is no more than

$$\mathbb{P}(\tilde{S}_2 \neq 0, \tilde{S}_m = 0) \leq \frac{(m-2)^2}{(2p)^2}, \quad (3.15)$$

where  $\tilde{S}_m$  denotes a simple random walk on  $\mathbb{Z}^p$  of length  $m \geq 3$ . The reason is that, conditioning on the event  $\{\tilde{S}_2 = x\}$  where  $x \neq 0$ , two of the next  $(m - 2)$  steps must go along the reverse directions of the coordinates used in the first two steps in order to return to the origin at the  $m$ -th step (note  $m$  must be even). This happens with probability no more than  $\frac{(m-2)^2}{(2p)^2}$ .

For all other values of  $m_1$  and  $m'_1$ , there are at least two bubble segments (e.g., Figure 5),

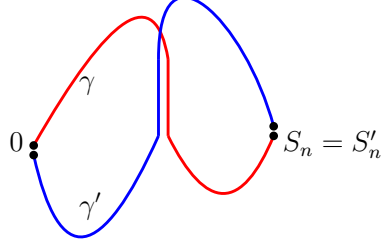


Figure 5: Representation of the case  $K = 1$  with a negative alignment of the overlapping segment. In this case two bubbles must be created. Parallel segments represent identical edges.

which we can easily estimate using (3.14). We use  $1 = K \leq l \leq n \leq C_0 \log p$  again and compute

$$2 \cdot \left(\frac{1}{2p}\right)^{l-1} \frac{(2n-2)^2}{(2p)^2} + \left(\frac{1}{2p}\right)^{l-1} \frac{(2n^2)^1}{(2p)^2} = \left(\frac{1}{2p}\right)^l o(p^{-1/2})$$

For  $K = 2$ , if  $\gamma$  or  $\gamma'$  has 2 bubble segments, we can use the strategy in Case (i) and (3.14). If there is only 1 bubble segments, then the two overlapping segments must attach to  $S_0 = S'_0 = 0$  and  $S_n = S'_n$ . Since  $\gamma$  and  $\gamma'$  are self-avoiding, this situation is quite similar to the fusion of bubbles in Figure 4 (a) and (b). The bubble in the middle can be easily estimated using (3.15). We leave the details to the readers.  $\square$

## 4 Proof of the Lower Bound

In this section we establish the desired lower bound. We assume the existence of a constant  $a \in [0, \infty)$  so that

$$\lim_{x \downarrow 0} \frac{\mathbb{P}(\tau_e \leq x)}{x} = a. \quad (4.1)$$

Note that this condition is weaker than (1.4).

**Proposition 4.1.** *Assume  $\mathbb{E}\tau_e < \infty$  and (4.1). Then,*

$$\liminf_{d \rightarrow \infty} \frac{\mu(e_1)ad}{\log d} \geq \frac{1}{2}.$$

For the proof of Proposition 4.1, we will need a few lemmas. Let  $b_n = T(0, H_n)$  be the passage time from the origin to the hyperplane  $H_n$ .

**Lemma 4.2.** *If for some constant  $x > 0$ ,*

$$\sum_n \mathbb{P}(b_n \leq xn) < \infty$$

*then  $\mu(e_1) \geq x$ .*

*Proof.* This is a consequence of the Borel-Cantelli lemma as the time constant  $\mu(e_1)$  is also the limit of  $b_n/n$  as  $n$  goes to infinity [7, Equation (1.13)].  $\square$

We now fix  $\delta > 0$  and set

$$x = \frac{(1 - \delta) \log d}{2ad}. \quad (4.2)$$

**Lemma 4.3.** *For any  $\delta > 0$ , there exists  $d$  large enough so that, for any fixed  $n \in \mathbb{N}$  and any  $k \geq n$ ,*

$$\mathbb{P}(S_k \leq nx) \leq \left( \frac{(1 + \delta)eanx}{k} \right)^k.$$

*Proof.* Fix  $\delta > 0$ . By (4.1), we can choose  $\varepsilon_0$  small enough so that

$$\mathbb{P}(\tau_e \leq t) \leq (1 + \delta/2)at < 1$$

for all  $t \in [0, \varepsilon_0]$ . Now, let  $Y$  be any nonnegative random variable with density  $f(y) = (1 + \delta/2)a$  on  $[0, \varepsilon_0]$ . For any  $0 \leq \varepsilon \leq \varepsilon_0$ , the random variable  $X_1 := \tau_e \mathbf{1}_{\{\tau_e < \varepsilon\}} + \varepsilon \mathbf{1}_{\{\tau_e \geq \varepsilon\}}$  stochastically dominates  $X_2 := Y \mathbf{1}_{\{Y < \varepsilon\}} + \varepsilon \mathbf{1}_{\{Y \geq \varepsilon\}}$  as for all  $t \in \mathbb{R}$ ,

$$\mathbb{P}(X_1 \leq t) \leq \mathbb{P}(X_2 \leq t).$$

Therefore for any non-increasing function  $\phi : \mathbb{R} \rightarrow \mathbb{R}$

$$\begin{aligned} \mathbb{E}\phi(\tau_e) \mathbf{1}_{\{\tau_e < \varepsilon\}} + \phi(\varepsilon) \mathbb{P}(\tau_e \geq \varepsilon) &\leq \mathbb{E}\phi(Y) \mathbf{1}_{\{Y < \varepsilon\}} + \phi(\varepsilon) \mathbb{P}(Y \geq \varepsilon) \\ &\leq \int_0^\varepsilon \phi(y) f(y) dy + \phi(\varepsilon). \end{aligned} \quad (4.3)$$

This implies that for any  $0 \leq \varepsilon \leq \varepsilon_0$  and any  $\gamma > 0$ , if we take  $\phi(t) = e^{-\gamma t}$  we have

$$\begin{aligned} \mathbb{P}(S_k \leq nx) &\leq e^{n\gamma x} \mathbb{E} e^{-\gamma S_k} \leq e^{n\gamma x} \left( \mathbb{E} e^{-\gamma \tau_e} \mathbf{1}_{\{\tau_e < \varepsilon\}} + e^{-\gamma \varepsilon} \mathbb{P}(\tau_e \geq \varepsilon) \right)^k \\ &\leq e^{n\gamma x} \left( \int_0^\varepsilon e^{-\gamma y} f(y) dy + e^{-\gamma \varepsilon} \right)^k \leq e^{n\gamma x} \left( \frac{(1 + \delta/2)a}{\gamma} + e^{-\gamma \varepsilon} \right)^k. \end{aligned}$$

Choose  $d$  large enough so that  $\sqrt{x} \leq \min\{\varepsilon_0, 1\}$  and such that for all  $y \geq 1$

$$y \exp\left(-\frac{y}{\sqrt{x}}\right) \leq \frac{\delta ax}{2}. \quad (4.4)$$

This is possible since the left side is monotonically decreasing in  $y$  on  $[1, \infty)$  if  $x \leq 1$ . Hence one suffices to find  $x \leq \min\{\varepsilon_0^2, 1\}$  such that

$$\exp\left(-\frac{1}{\sqrt{x}}\right) \leq \frac{\delta ax}{2} \Leftrightarrow \sqrt{x} \log \frac{\delta a}{2} + 2\sqrt{x} \log \sqrt{x} \geq -1, \quad (4.5)$$

which is possible by choosing  $d$  large enough and making each term above greater than  $-1/3$ . Now set  $\gamma = k/(nx)$ . By taking

$$\varepsilon = \frac{k}{\gamma n \sqrt{x}} = \sqrt{x} \in [0, \varepsilon_0], \quad y = \frac{k}{n} \geq 1,$$



we have by (4.4)

$$\mathbb{P}(S_k \leq nx) \leq \left( \frac{eax(1 + \delta/2)}{y} + e^{-\frac{y}{\sqrt{x}}+1} \right)^k \leq \left( \frac{eax(1 + \delta)}{y} \right)^k, \quad (4.6)$$

which proves the Lemma.  $\square$

We still need one combinatorial estimate that we take from [7, (6.20)].

**Lemma 4.4.** *The number  $\mathcal{N}_{k,n}$  of lattice paths in  $\mathbb{Z}^d$  from 0 to  $H_n$  of  $k$  steps is at most*

$$(2d)^k \min \left( 1, \exp \left( -n\rho + \frac{k}{d}(\cosh \rho - 1) \right) \right)$$

for any  $\rho \geq 0$ .

*Proof of Proposition 4.1.* For  $0 < \delta < 1$  fixed choose  $x$  as in (4.2).

We will use the union bound

$$\mathbb{P}(b_n \leq nx) \leq \sum_{k=n}^{\infty} \mathcal{N}_{k,n} \mathbb{P}(S_k \leq nx).$$

Set  $M = 4enaxd = 2en(1 - \delta) \log d$ . Using Lemma 4.4 with  $\rho = \log(\frac{2dn}{k})$  for  $n \leq k \leq M$  and Lemma 4.3 we have

$$\mathbb{P}(b_n \leq nx) \leq \sum_{n \leq k \leq M} \left( \frac{ke}{2dn} \right)^n \left( \frac{2d(1 + \delta)eanx}{k} \right)^k + \sum_{k > M} \left( \frac{2d(1 + \delta)eanx}{k} \right)^k \quad (4.7)$$

Choose  $\delta \leq 1/2$ , the second sum in the right side of (4.7) is bounded above by

$$\left( \frac{1 + \delta}{2} \right)^M \frac{1}{1 - \frac{1 + \delta}{2}} \leq 4 \left( \frac{3}{4} \right)^{2en(1 - \delta) \log d}$$

which is summable in  $n$ . On the other hand, if we write  $z = k/n$ , a little algebra implies that the first sum in (4.7) is bounded above by

$$M \left[ \frac{e}{2d} \max_z \left( z^{1-z} [(1 - \delta^2)e \log d]^z : 1 \leq z \leq 2e(1 - \delta) \log d \right) \right]^n. \quad (4.8)$$

The term inside the large square bracket above is bounded by

$$2e^2(1 - \delta) \frac{\log d}{2d} \max_z \left( \left[ \frac{(1 - \delta^2)e \log d}{z} \right]^z : 1 \leq z \leq 2e(1 - \delta) \log d \right) \quad (4.9)$$

As for any  $c > 0$ , the function  $f(z) = (c/z)^z$  has a maximum value equal to  $e^{c/e}$  on  $0 \leq z \leq c$ , we obtain that (4.9) is bounded above by

$$e^2(1 - \delta) \frac{\log d}{d} \exp((1 - \delta^2) \log d) \leq e^2 d^{-\delta^2} \log d. \quad (4.10)$$

Given the choice of  $\delta$ , this term is strictly bounded above by 1 for  $d$  large enough, which turns (4.8) summable in  $n$ . Now the proposition follows from Lemma 4.2 with  $x = \frac{(1 - \delta) \log d}{2ad}$  and sending  $\delta$  to 0.  $\square$

We now end this subsection with the proof of Theorem 1.2 in the case  $a = \infty$  and  $a = 0$ . In this case, assumption (1.5) implies that for any  $M > 0$  it is possible to find  $\varepsilon > 0$  such that for any  $x < \varepsilon$ ,  $\mathbb{P}(\tau_e \leq x) \geq xM$ . Let  $Y_M$  be a random variable with density  $f(y) = My$  on  $[0, \varepsilon]$  and such that for any  $t \in \mathbb{R}$

$$\mathbb{P}(Y_M \leq t) \leq \mathbb{P}(\tau_e \leq t).$$

As  $\mathbb{P}(Y_M \leq \varepsilon) \leq \mathbb{P}(\tau_e \leq \varepsilon)$ , this random variable can be constructed by simply choosing a non-decreasing function on  $[\varepsilon, \infty)$  that has limit 1 at  $\infty$  and is bounded above by  $\mathbb{P}(\tau_e \leq t)$ . This way, we can use the comparison theorem of van den Berg-Kesten [10, Theorem 2.13] to obtain for any  $d \geq 2$ ,  $\mu(e_1) \leq \mu^{Y_M}(e_1)$ , where  $\mu^{Y_M}$  is the time constant for FPP in  $\mathbb{Z}^d$  with passage times distributed according to  $Y_M$ . Since  $Y_M$  satisfies the hypothesis of Theorem 1.2 with  $a = M$ , we get

$$\limsup_{d \rightarrow \infty} \frac{\mu(e_1)d}{\log d} \leq \lim_{d \rightarrow \infty} \frac{\mu^{Y_M}(e_1)d}{\log d} = \frac{1}{2M}.$$

Taking  $M$  to infinity gives us the desired result.

The case  $a = 0$  is similar. For any  $m > 0$  we construct a random variable  $Y_m$  that dominates  $\tau_e$  as  $\mathbb{P}(Y_m \leq t) \geq \mathbb{P}(\tau_e \leq t)$  and satisfies (1.4) with  $a = m$ . Van den Berg-Kesten comparison theorem combined with the result for  $a > 0$  implies

$$\liminf_{d \rightarrow \infty} \frac{\mu(e_1)d}{\log d} \geq \lim_{d \rightarrow \infty} \frac{\mu^{Y_m}(e_1)d}{\log d} = \frac{1}{2m}.$$

The result follows by taking  $m$  to zero.

## 5 Application to the limit shape

In this section, we will exclude certain candidates of possible limit shapes in high dimension, including the Euclidean ball. The method here is the same as the one used by Kesten [7]. We will compare the time constant  $\mu(e_1)$  in the  $e_1$  direction with time constant  $\mu^*$  in the diagonal direction. Here  $\mu^*$  is defined as

$$\mu^* := \lim_{n \rightarrow \infty} \frac{T(\mathcal{J}_n)}{n}, \quad \text{a.s.},$$

where  $\mathcal{J}_n$  is the hyperplane defined by  $\mathcal{J}_n := \{(x_1, \dots, x_d) : x_1 + x_2 + \dots + x_d = n\sqrt{d}\}$  and the limit hold in  $L^1$ . Without any loss in what follows, we slightly abuse our notation by taking  $\sqrt{d}$  as the smallest integer greater than the square root of  $d$ . It has been shown in [2] that, when  $\tau_e$  follows a standard exponential distribution, then for  $d \geq 2$ ,

$$\mu^* \geq \frac{\sqrt{\alpha_*^2 - 1}}{2\sqrt{d}} \geq \frac{0.3313}{\sqrt{d}}, \quad (5.1)$$

where  $\alpha_*$  is the non null solution of  $\coth \alpha = \alpha$ . Recently, still under the assumption of exponential passage times, Martinsson [9] proved a matching upper bound establishing

$$\lim_{d \rightarrow \infty} \sqrt{d}\mu^* = \frac{\sqrt{\alpha_*^2 - 1}}{2}.$$

## 5.1 A lower bound in the diagonal direction

We start by showing that the lower bound (5.1) is also true under our setting in the large  $d$  limit:

**Theorem 5.1.** *Suppose the edge weight distribution satisfies (1.4) and  $\mu^* = \mu_d^*$  is the time constant in the diagonal direction as defined above. Then*

$$\liminf_{d \rightarrow \infty} \sqrt{d} \mu^* \geq \frac{\sqrt{\alpha_*^2 - 1}}{2a},$$

*Remark 5.2.* The result also holds under (4.1), which is more general than (1.4) and the proof is a slight modification of that in Section 4 and [2]. However, we state Theorem 5.1 and provide its proof under (1.4) for two purposes: (i) to quantify the difference between  $\mu^*$  and  $\frac{\sqrt{\alpha_*^2 - 1}}{2a\sqrt{d}}$ , for  $d$  large, and (ii) to compare the lower bound of  $\mu^*$  with the upper bound of  $\mu$  for finite but large  $d$ , under the same set of conditions.

*Proof.* For  $\delta \in (0, 1)$ , we may always choose  $d$  large enough such that

$$d \geq \frac{\alpha_*^2 - 1}{4a^2} \max \{ \varepsilon_0^{-4}, e^{8C/\delta}, 1 \}. \quad (5.2)$$

Recall that  $\varepsilon_0$  is the right endpoint of the interval  $[0, \varepsilon_0]$  on which the distribution of  $\tau_e$  satisfies (1.4) with constants  $a$  and  $C$ . We set  $x = \frac{(1-\delta)\sqrt{\alpha_*^2 - 1}}{2a\sqrt{d}}$ . Note that  $\sqrt{x} \in [0, \min\{\varepsilon_0, 1\}]$  due to (5.2). Then for any fixed  $n \in \mathbb{N}$  and any  $k \geq n\sqrt{d}$ , we can repeat the computation before (4.4) to obtain

$$\mathbb{P}(S_k \leq nx) \leq \left( \frac{(1+\delta)eanx}{k} \right)^k. \quad (5.3)$$

To see this, one can just follow the proof of Lemma 4.3 with  $y = \frac{k}{n\sqrt{d}}$  and  $\gamma = k/(nx\sqrt{d})$ , provided (4.5) holds for our choice of  $x$  and any  $y \geq 1$ , i.e.,

$$\frac{\sqrt{2a}}{(\alpha_*^2 - 1)^{1/4}} d^{1/4} - \frac{1}{2} \log d \geq \log \frac{4}{\delta(1-\delta)\sqrt{\alpha_*^2 - 1}}. \quad (5.4)$$

This is always possible by choosing  $d$  large enough. Next, we use the upper bound for the number  $D_k^{(n)}$  of self-avoiding walks of length  $k$  from 0 to  $\mathcal{J}_n$  from Lemma 3 of [2]:

$$\begin{aligned} D_k^{(n)} &\leq (2d)^k \frac{n\sqrt{d}}{k} \binom{k}{(k+n\sqrt{d})/2} 2^{-k} \\ &\leq \sqrt{\frac{1}{n\sqrt{d}}} \left( \frac{2dy}{(y+1)^{(y+1)/(2y)}(y-1)^{(y-1)/2y}} \right)^k. \end{aligned}$$

Hence, we have

$$\begin{aligned}
\mathbb{P}(T(\mathcal{J}_n) \leq nx) &\leq \sum_{k \geq n\sqrt{d}} D_k^{(n)} \mathbb{P}(S_k \leq nx) \\
&\leq \sqrt{\frac{1}{n\sqrt{d}}} \sum_{k \geq n\sqrt{d}} \left( \frac{2\sqrt{d}(1+\delta)eax}{(y+1)^{(y+1)/(2y)}(y-1)^{(y-1)/2y}} \right)^k \\
&\leq \sqrt{\frac{1}{n\sqrt{d}}} \sum_{k \geq n\sqrt{d}} \left( \frac{(1-\delta^2)e\sqrt{\alpha_*^2-1}}{\inf_{y \geq 1} \{(y+1)^{(y+1)/(2y)}(y-1)^{(y-1)/2y}\}} \right)^k.
\end{aligned}$$

Note that  $\inf_{y \geq 1} \{(y+1)^{(y+1)/(2y)}(y-1)^{(y-1)/2y}\} = e\sqrt{\alpha_*^2-1}$ , and hence  $\mathbb{P}(T(\mathcal{J}_n) \leq nx)$  decays exponentially. The conclusion follows from the Borel-Cantelli lemma.  $\square$

The proof above implies the following quantitative lower bound.

**Corollary 5.3.** *The bound*

$$\mu^* \geq \frac{\sqrt{\alpha_*^2-1}}{2a\sqrt{d}}(1-\delta) \quad (5.5)$$

holds for all  $\delta \in (0, 1)$  and  $d \geq 1$  such that (5.2) and (5.4) are satisfied.

*Proof of Theorem 1.5.* Firstly, choose  $d_1$  sufficiently large such that (5.2) and (5.4) are satisfied with  $\delta = 1/2$ , for all  $d > d_1$ . This means  $\mu^* \geq \frac{1}{2} \frac{\sqrt{\alpha_*^2-1}}{2a\sqrt{d}}$  for all  $d > d_1$ . Secondly, due to Theorem 1.2, we can choose  $d_2$  large enough such that  $\mu(e_1) \leq \frac{\log d}{ad}$  for all  $d > d_2$ . Choose  $d_3$  large enough such that  $\frac{1}{2} \frac{\sqrt{\alpha_*^2-1}}{2a\sqrt{d}} > \frac{\log d}{ad}$  for all  $d > d_3$ . Putting  $d_0 = \max(d_1, d_2, d_3)$ , we have for all  $d > d_0$ ,

$$\mu(e_1) \leq \frac{\log d}{ad} < \frac{1}{2} \frac{\sqrt{\alpha_*^2-1}}{2a\sqrt{d}} \leq \mu^*,$$

which means the intersection of  $\mathcal{B}$  with the line  $\ell_d = \{\lambda(1, 1, \dots, 1) : \lambda \in \mathbb{R}\}$  is strictly contained in the Euclidean ball  $\mathbf{B}$ . This shows that  $\mathcal{B} \neq \mathbf{B}$  and  $\mathcal{B} \subsetneq \mathbf{C}$ .

Note that to show that the limit shape is not equal to the  $d$ -dimensional diamond  $\mathbf{D}$ , it suffices to show that

$$\mu^* < \sqrt{d}\mu(e_1). \quad (5.6)$$

The term  $\sqrt{d}$  appears due to our choice of normalization of the hyperplane  $\mathcal{J}_n$ . To prove (5.6), construct a path  $\gamma^n = \{\langle v_i, v_{i+1} \rangle\}_{i=1, \dots, n\sqrt{d}}$  from 0 to  $\mathcal{J}_n$  such that

- (a) For every  $i = 1, \dots, n\sqrt{d}$ , the  $i$ -th edge  $\langle v_i, v_{i+1} \rangle$  of  $\gamma^n$  is always in the positive direction, that is,  $v_{i+1} - v_i \in \{e_1, \dots, e_d\}$ .
- (b)  $\langle v_i, v_{i+1} \rangle$  is chosen as the edge with the smallest weight coming out of  $v_i$ , that is,

$$\tau_{\langle v_i, v_{i+1} \rangle} = \min_{k=1, \dots, d} \{\tau_{\langle v_i, v_i + e_k \rangle}\}.$$

If we let

$$Y = \min \{t_1, \dots, t_d\},$$

where the  $t_i$ 's are independent copies of  $\tau_e$ , we see by construction,

$$\mathbb{E}T(\gamma^n) = n\sqrt{d}\mathbb{E}Y. \quad (5.7)$$

Hence,

$$\mu^* = \lim_{n \rightarrow \infty} \frac{\mathbb{E}T(\mathcal{J}_n)}{n} \leq \lim_{n \rightarrow \infty} \frac{\mathbb{E}T(\gamma^n)}{n} = \sqrt{d}\mathbb{E}Y.$$

Our assumptions on the distribution of  $\tau_e$  imply the existence of a positive constant  $c$ , independent of  $d$ , such that  $\mathbb{E}Y \leq c/d$  for all  $d \geq 2$ . Indeed, by hypothesis (1.4) and the fact that  $a > 0$ , we can find  $\delta, \varepsilon > 0$  so that

$$\mathbb{P}(\tau_e > x) = 1 - \mathbb{P}(\tau_e \leq x) \leq 1 - xa(1 - \varepsilon) < 1$$

for all  $x \in [0, \delta]$  and such that  $\mathbb{P}(\tau < \delta) < 1$ . Choosing a constant  $m > \mathbb{E}\tau_e$ , we write

$$\begin{aligned} \mathbb{E}Y &= \int_0^\infty \mathbb{P}(\tau_e > t)^d dt \leq \int_0^\delta (1 - at(1 - \varepsilon))^d dt + \int_\delta^m \mathbb{P}(\tau_e > t)^d dt + \int_m^\infty \mathbb{P}(\tau_e > t)^d dt \\ &\leq \frac{1}{d+1} \frac{1}{a(1 - \varepsilon)} + (m - \delta)\mathbb{P}(\tau_e > \delta)^d + \left(\frac{\mathbb{E}\tau_e}{m}\right)^d \frac{m}{d-1}, \end{aligned} \quad (5.8)$$

where in the first integral we used hypothesis (1.4) and in the last one we used Markov's inequality. As  $\mathbb{P}(\tau_e > \delta) < 1$ , we can find  $c > 0$  such that all three terms are bounded above by  $c/d$ , for all  $d \geq 2$ . Combining this with Proposition 4.1, we have for any  $d$  large enough

$$\mu^* \leq \frac{c}{\sqrt{d}} \leq \frac{\log d}{3a\sqrt{d}} < \sqrt{d}\mu(e_1),$$

proving (5.6). □

## Appendix

In this section, we show how to compute a quantitative upper bound of  $\mu(e_1)$ . This bound will only depend on  $d$ ,  $\mathbb{E}\tau_e$  and  $a, \varepsilon_0, C$  satisfying (1.4).

We proceed as follows. First, slightly modifying (3.4) in Section 3, we notice that for any  $A > 0, \delta \in (0, 1)$  and  $B \neq 0$  satisfying

$$\mathbb{E}(N_{n,x}^2) \leq A(\mathbb{E}N_{n,x})^2 \quad \text{and} \quad y := \frac{B\delta \log d}{2ad} \leq \min\{\varepsilon_0, e^{-C}\}, \quad (5.9)$$

and any  $\eta > 0$ ,

$$\mu(e_1) \leq \frac{\log d}{2ad} \left( B\delta + \frac{1}{1 - \delta} \right) + \left( 1 - \frac{B\delta \log d}{2Ad} \left[ 1 - \frac{C}{|\log y|} \right] \right)^{\left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor} \mathbb{E}\tau_e. \quad (5.10)$$

Letting  $\Upsilon(A, B, \delta, \eta)$  be the right side of (5.10), we obtain

$$\mu(e_1) \leq \inf_{(\delta, A, B, \eta)} \Upsilon(A, B, \delta, \eta), \quad (5.11)$$

where the infimum is taken over all  $\eta > 0$  and  $A, B, \delta$  satisfying (5.9).

As we will see, the main task here is to find any (but preferably the smallest)  $A$  that satisfies (5.9). In Section 3.4, we found such a number for any given  $\delta \in (0, 1)$  and  $\eta > 0$ . Indeed, we saw that we can choose any  $A = A(\delta, \eta)$  such that

$$A \geq 1 + e^{f_{a,C}(\delta, d)g_\eta^2(\delta, d)} (1 + o(p^{-1/2})) \sum_{l=0}^{n-1} \left( \frac{1 - \delta}{1 - \frac{1}{\delta^{1+\eta} \log d}} \right)^l + \frac{1}{\mathbb{E}N_{n,x}}, \quad (5.12)$$

where upper bounds of  $1/\mathbb{E}N_{n,x}$ ,  $f_{a,C}(\delta, d)$ , and  $g_\eta(\delta, d)$  can be recovered from (3.2), (3.8) and (3.9), respectively, once we are given  $\delta, \eta, d$  and the parameters  $a, C$  in (1.4) for the distribution of  $\tau_e$ . The issue now is to control the  $o(p^{-1/2})$  term, where  $p = d - \lfloor \frac{d}{\delta^{1+\eta} \log d} \rfloor > 0$ . This will be done in the next section of the appendix. In the last section, we provide a few specific examples of these computations.

## A refinement of Lemma 3.4

To get a quantitative estimate of the lower bound of  $A$  using (5.12), we need to get control of the  $o(p^{-1/2})$  term. This term comes exactly from the estimate on the probabilities in Lemma 3.4. We will control these probabilities by computing the combinatorics in the “bubble” argument more precisely. We start from the two cases defined in (3.11) and (3.12).

For the first case, it is not difficult to see from the proof of Lemma 3.4(i) that,

$$\mathbb{P}(\text{Case (i)}) \leq \left( \frac{1}{2p} \right)^l \left\{ 1 + \frac{(n-l-1)^2}{p} + 2p \sum_{K=2}^l \frac{\binom{l-1}{K-1} \left[ \binom{n-1}{K} \cdot K! \right]^2}{p^K} \right\} = \left( \frac{1}{2p} \right)^l [1 + (\text{I})],$$

where the right side is obtained by replacing  $n$  in Lemma 3.4(i) by  $(n-1)$  and following the combinatorics there.

For the second case, we provide more detail here. As before, we let  $\tilde{\gamma}$  and  $\tilde{\gamma}'$  be the paths obtained by removing the last step of  $\gamma$  and  $\gamma'$ , respectively. We use  $B_K(\tilde{\gamma}, \tilde{\gamma}')$  to denote the event that the overlapping edges in  $\tilde{\gamma}$  and  $\tilde{\gamma}'$  are clustered in  $K$  segments. We compute  $K = 1, 2, 3$  case separately, whereas all  $K \geq 4$  cases are considered together.

When  $K = 1$  and the overlapping segments in  $\gamma$  and  $\gamma'$  align in the same direction, there can be either one bubble or two bubbles. The “one-bubble” diagram is in Figure 4. Hence

$$\begin{aligned} \mathbb{P}(\text{Case (ii)}; B_1(\tilde{\gamma}, \tilde{\gamma}')) &\leq 2 \cdot \left( \frac{1}{2p} \right)^{l-1} \left( \frac{2(n-l)-2}{2p} \right)^2 \\ &\quad + \left( \frac{1}{2p} \right)^{l-1} \sum_{m_1, m'_1=1}^{n-l-1} \left( \frac{m_1 + m'_1 - 2}{2p} \right)^2 \left( \frac{2n-2l-m_1-m'_1-2}{2p} \right)^2 \\ &\quad + \left( \frac{1}{2p} \right)^{l-1} (n-l-1)^2 \left( \frac{1}{2p} \right)^2 = \left( \frac{1}{2p} \right)^l \cdot (\text{II}) \end{aligned}$$

When  $K = 2$  and there is only one bubble, both overlapping segments have to align in the same direction and attach to the start and endpoint of  $\tilde{\gamma}$  and  $\tilde{\gamma}'$ . This diagram occurs with probability no more than

$$(l-2) \left( \frac{1}{2p} \right)^{l-1} \left( \frac{2n-2l-2}{2p} \right)^2,$$

where  $(l-2)$  is the number of ways of grouping the  $(l-1)$  overlapping edges into two consecutive overlapping segments,  $\left( \frac{2n-2l-2}{2p} \right)^2$  is the factor for the middle bubble, applying (3.15). We can also estimate the probabilities for other types of alignments for  $K = 2$  case, which gives

$$\begin{aligned} \mathbb{P}(\text{Case (ii)}; B_2(\tilde{\gamma}, \tilde{\gamma}')) &\leq \left( \frac{1}{2p} \right)^{l-1} (l-2) \left[ 6 \cdot \left( \frac{n-l-1}{2p} \right)^2 + \right. \\ &\quad \left. \frac{2(n-l-1)^2(n-l-2)^4}{(2p)^4} + \frac{4(n-1)^2(n-l)^2}{(2p)^3} \right] \\ &= \left( \frac{1}{2p} \right)^l \cdot (\text{III}). \end{aligned}$$

Following the same strategy, we can compute the probabilities for the  $K = 3$  case

$$\begin{aligned} \mathbb{P}(\text{Case (ii)}; B_3(\tilde{\gamma}, \tilde{\gamma}')) &\leq \binom{l-2}{2} \left( \frac{1}{2p} \right)^{l-1} \left[ \frac{2(n-l)^2}{(2p)^2} + \frac{8(n-1)^4(n-l)^2}{(2p)^3} \right] \\ &= \left( \frac{1}{2p} \right)^l \cdot (\text{IV}), \end{aligned}$$

and the union of all  $K \geq 4$  case

$$\sum_{K=4}^l \mathbb{P}(\text{Case (ii)}; B_K(\tilde{\gamma}, \tilde{\gamma}')) \leq \left( \frac{1}{2p} \right)^{l-1} \cdot 2p \sum_{K=4}^l \frac{\binom{l-2}{K-1} \left[ \binom{n-1}{K} \cdot K! \right]^2}{p^K} = \left( \frac{1}{2p} \right)^l \cdot (\text{V}).$$

Combining everything above gives a quantitative upper bound for  $\mathbb{P}(\text{Case (ii)})$ . The quantitative upper bound can be reduced even further by computing  $\mathbb{P}(\text{Case (ii)}; B_K(\tilde{\gamma}, \tilde{\gamma}'))$  separately for more  $K$ 's, but we stop here.

Therefore,  $o(p^{-1/2})$  is no more than summing (I) through (V).

## Special Case: Exponential Distribution

**Limit shape is not a  $d$ -dimensional cube for  $d \geq 269,000$ .**

For the special case of  $\tau_e$  following a standard exponential distribution, we have  $a = 1, C = o(1)$ , and  $\mathbb{E}\tau_e = 1$ . Moreover,  $\mu^* \geq \frac{\sqrt{\alpha_*^2 - 1}}{2\sqrt{d}}$  for all  $d \geq 2$ . To see this, one can either check [2] or go through the proof of Theorem 5.1, using the fact that (5.3) holds for  $\delta = 0$ .

To estimate  $\mu(e_1)$ , instead of using Lemma 3.2, we can compute  $\mathbb{P}(S_n \leq x)$  explicitly in this case,

$$\mathbb{P}(S_n \leq x) = \int_0^x \frac{y^{n-1} e^{-y}}{(n-1)!} dy := \gamma(n, x)$$

where  $\gamma(n, x)$  is the cumulative distribution function of a gamma distribution with shape parameter  $n$  and scale parameter 1. This allows us to replace the ratio  $\frac{\mathbb{P}(S_{n-l} \leq x)}{\mathbb{P}(S_n \leq x)}$  by  $\frac{\gamma(n-l, x)}{\gamma(n, x)}$ . We also use  $\gamma(n, x)$  to estimate  $1/\mathbb{E}N_{n,x}$

$$\frac{1}{\mathbb{E}N_{n,x}} = \frac{1}{|\mathcal{P}_n| \cdot \mathbb{P}(S_n \leq x)} \leq \frac{1}{[2p-1-\log(2p-1)]^{n-1} \gamma(n, x)}.$$

Therefore, following exactly the same computation in Section 3, we obtain that we can choose any

$$A \geq 1 + \frac{1}{\mathbb{E}N_{n,x}} + e^{2g_\eta(\delta, d)} \sum_{l=1}^{n-1} \frac{\gamma(n-l, x)}{\gamma(n, x)} [\mathbb{P}(\text{Case (i)}) + \mathbb{P}(\text{Case (ii)})], \quad (5.13)$$

to obtain

$$\mu(e_1) \leq \inf_{(\delta, B, \eta)} \left\{ \frac{\log d}{2d} \left( B\delta + \frac{1}{1-\delta} \right) + \left( 1 - \frac{1 - e^{-\frac{B\delta \log d}{2d}}}{A} \right)^{\left\lfloor \frac{d}{\delta^{1+\eta} \log d} \right\rfloor} \right\}. \quad (5.14)$$

For instance, plugging in  $d = 268,337$ ,  $\eta = 10^{-3}$ ,  $\delta = 0.764$  and  $B = 23.85$ , we can take  $A = 1.20$  and  $\mu(e_1) \leq 0.000638$  and  $\mu^* \geq 0.000639 > \mu(e_1)$ . We checked that for all  $d \in (268337, 10^6)$ ,  $\frac{\sqrt{\alpha_*^2 - 1}}{2\sqrt{d}}$  is greater than the right side of (5.14). The case  $d \geq 10^6$  follows directly from Kesten [7, Remark 8.5].

### Limit shape is not a $d$ -dimensional diamond for any $d \geq 110$ .

Recall that  $Y := \min\{t_1, \dots, t_d\}$ , where  $t_i$ 's are i.i.d. copies of  $\tau_e$ . For the standard exponential distribution we have  $\mathbb{E}Y = d^{-1}$  in (5.7) so  $\mu^* \leq d^{-1/2}$  for any  $d \geq 2$ . At the same time, we can take  $\delta = 0$  in Lemma 4.3, which implies that the exponent of  $\delta$  in (4.8) changes from 2 to 1. This leads that (4.7) and (4.10) are summable for any  $\delta \in (0, 1)$  and all  $d$  such that

$$e^2 d^{-\delta} (1-\delta) \log d < 1, \quad (5.15)$$

which gives, by Lemma 4.2, the bound  $\sqrt{d}\mu(e_1) \geq \frac{(1-\delta)\log d}{2\sqrt{d}}$  for this choice of  $\delta$  and  $d$ . This implies that the limit shape is not a  $d$ -dimensional cube (as  $\mu^* < \sqrt{d}\mu(e_1)$ ) for any  $d$  that satisfies

$$d > \exp\left(\frac{2}{1-\delta}\right)$$

and (5.15). Choosing  $\delta = (2 + \log 2)/(4 + \log 2)$  we see that the equations above are satisfied for any  $d \geq 110$ .

If  $\tau_e$  is not exponentially distributed and we only have  $\mathbb{E}Y \leq d^{-1}$ , for instance, when  $\tau_e \sim \mathcal{U}[0, 1]$ , we must keep the exponent of  $\delta$  in (4.8) equal to 2 and the choice of  $\delta = 0.669$  excludes the  $d$ -dimensional diamond for  $d \geq 416$ .



## References

- [1] Antonio Auffinger, Michael Damron, and Jack Hanson, *50 years of first passage percolation*, arXiv.org (2015), 1–148.
- [2] Olivier Couvonné, Nathanaël Enriquez, and Lucas Gerin, *Construction of a short path in high-dimensional first passage percolation*, Electron. Commun. Probab. **16** (2011), 22–28. MR 2753301 (2012e:60254)
- [3] J. Theodore Cox and Richard Durrett, *Oriented percolation in dimensions  $d \geq 4$ : bounds and asymptotic formulas*, Math. Proc. Cambridge Philos. Soc. **93** (1983), no. 1, 151–162. MR 684285 (84e:60150)
- [4] Deepak Dhar, *First passage percolation in many dimensions*, Phys. Lett. A **130** (1988), no. 4-5, 308–310. MR 949856 (89d:82083)
- [5] John Michael Hammersley, *Long-Chain Polymers and Self-Avoiding Random Walks*, Sankhyā: The Indian Journal of Statistics, Series A **25** (1963), no. 1, 29–38.
- [6] Harry Kesten, *On the number of Self-Avoiding Walks. II*, Journal of Mathematical Physics **5** (1964), no. 8, 1128–1137.
- [7] ———, *Aspects of first passage percolation*, École d’été de probabilités de Saint-Flour, XIV—1984, Lecture Notes in Math., vol. 1180, Springer, Berlin, 1986, pp. 125–264. MR 876084 (88h:60201)
- [8] Neal Madras and Gordon Slade, *The Self-Avoiding Walk*, Springer Science & Business Media, Boston, MA, November 2013.
- [9] Anders Martinsson, *First-passage percolation on cartesian power graphs*, arXiv.org (2015), 1–27.
- [10] Jacob van den Berg and Harry Kesten, *Inequalities for the time constant in first-passage percolation*, Ann. Appl. Probab. **3** (1993), no. 1, 56–80. MR 1202515 (94a:60134)